For since the fabric of the Universe is most perfect and the works of a most wise creator, nothing at all takes place in the Universe in which some rule of maximum or minimum does not appear

- Leonhard Euler

Chapter 3

Normalized Scale-Space Derivatives: A Statistical Analysis

This chapter presents a statistical analysis of multiscale derivative measurements. Noisy images and multiscale derivative measurements made of noisy images are analyzed; the means and variances of the measured noisy derivatives are calculated in terms of the parameters of the probability distribution function of the initial noise function and the scale or sampling aperture. Normalized and unnormalized forms of differential scale space are analyzed, and the statistical results are compared. A discussion of the results and their ramifications for multiscale analysis is included.

3.1. Introduction and Background

There has been substantial research in the area of differential invariants of scale space. Notably, researchers such as Koenderink, ter Haar Romeny, Florack, Lindeberg, Blom, and Eberly have contributed many papers on scale space and the invariances of scale-space derivatives [Koenderink 1984, ter Haar Romeny 1991ab, Florack 1993, Lindeberg 1992, Blom 1993, and Eberly 1994]. This sub-field of computer vision has yielded many significant insights. This chapter is an exploration of some of the statistical aspects of scale space when the source images are subject to spatially uncorrelated noise.

This section amplifies scale space and multiscale derivative concepts presented in Chapter 2. It also presents a notation for multiscale derivatives borrowed from other related research.

Given some scale or measurement aperture σ, scale-space derivatives of a digital image are measured by convolving the image with a derivative-of-Gaussian kernel. Given a continuous 2D image function $I(\mathbf{p})$ and a Gaussian kernel $G(\sigma, \mathbf{p})$ where $\mathbf{p} = (x, y)$, multiscale derivatives of arbitrary order are described by the following equation:

$$L_{x^n y^m} = L_{I, x^n y^m}(\mathbf{p} \mid \sigma) = \frac{\partial^n}{\partial x^n}\frac{\partial^m}{\partial y^m} I(\mathbf{p} \mid \sigma) = \frac{\partial^n}{\partial x^n}\frac{\partial^m}{\partial y^m} G(\sigma, \mathbf{p}) \otimes I(\mathbf{p}) \qquad (3\text{-}1)$$

The term shown above is described as *the n-th derivative in the x-direction and the m-th derivative in the y-direction given scale* σ. Ter Haar Romeny *et al.* have adapted Einstein

Summation Notation (ESN) to provide a compact description of scale-space derivatives. In their notation, shown in the leftmost expression in equation (3-1), the image function $I(\mathbf{p})$, the scale parameter σ, and the location parameter $\mathbf{p}$ are often assumed, resulting in the abbreviated notation $L_{x^n y^m}$ to represent scale-space differentiation. I use the abbreviated notation, $L_{x^n y^m}$, as well as the term $L_{I,x^n y^m}(\mathbf{p} \mid \sigma)$ interchangeably in this chapter. When it is necessary to specify a particular image function, spatial location, or scale, I will use the notation $L_{I,x^n y^m}(\mathbf{p} \mid \sigma)$.

The geometry captured in derivative measurements can be used to classify points within an image. Derivatives embody properties such as Gaussian or mean intensity curvature, gradient magnitude (often used as a measure of boundariness), isophote curvature, and a host of other important features of the image.

Derivatives can be measured not only along the directions of the Cartesian coordinate directions, but in any arbitrary direction about a point. The work of ter Haar Romeny and Florack describes a coordinate frame or gauge at each spatial location based on the structure of first order derivatives. When measured in gauge coordinates, higher order derivatives exhibit invariance with respect to spatial rotation. Expressing derivatives in gauge coordinates simplifies the notation and computation of scale-space derivatives [ter Haar Romeny 1991a].

Tracing image intensities and derivative values through changing scale is also often useful. Derivatives with respect to scale are easily computed. The Gaussian has the property that the second derivative with respect to spatial coordinates is directly proportional to the first derivative with respect to scale. This property simplifies many scale-space calculations. Multiscale derivatives of the continuous image function $I(\mathbf{p})$ given a Gaussian scale operator $G(\sigma, \mathbf{p})$ taken with respect to scale σ are

$$L_{\sigma^k} = L_{I,\sigma^k}(\mathbf{p} \mid \sigma) = \frac{\partial^k}{\partial \sigma^k} I(\mathbf{p} \mid \sigma) = \frac{\partial^k}{\partial \sigma^k} G(\sigma, \mathbf{p}) \otimes I(\mathbf{p}) \tag{3-2}$$

3.1.1. Scale-Space Differential Invariants

As described in Chapter 2, the Gaussian is a natural scale-space aperture function by the *a priori* constraints of linearity, shift invariance, and rotation invariance. Zoom invariance is achieved by imposing an appropriate metric on scale-space. Florack and ter Haar Romeny specified a distance metric that preserves the Euclidean nature of scale space [ter Haar Romeny 1991ab]. In more recent work, Eberly suggests a dimensionless 1-form to be used in scale-space measurements, specifying a hyperbolic construction for scale space [Eberly 1994ab].

$$\frac{\partial \mathbf{p}}{\sigma} \qquad \rho \frac{\partial \sigma}{\sigma} \tag{3-3}$$

ρ is a constant relating rate of change in the scale dimension to spatial rate of change. In most uses of this 1-form, $\rho = 1$.

Use of this 1-form suggests that to exhibit zoom-invariance, scale-space derivatives must be normalized by the scale of the differential operators. For example, in Fritsch's study of the multiscale medial axis (now called *core*), he applied an operator with its kernel equal to the Laplacian of Gaussian multiplied by σ^2 as a filter for detecting medialness. He showed that this operator exhibited zoom invariance [Fritsch 1993].

Eberly's research on scale-space derivatives generalizes the normalization process for derivatives of arbitrary order. The form of a normalized scale-space spatial derivative $\hat{L}_{x^n y^m}$ is

$$\hat{L}_{x^n y^m} = \hat{L}_{I, x^n y^m}(\mathbf{p} \mid \sigma) = \sigma^{n+m}\left(\frac{\partial^n}{\partial x^n}\frac{\partial^m}{\partial y^m} G(\sigma, \mathbf{p}) \otimes I(\mathbf{p})\right) \tag{3-4}$$

The multiscale derivatives with respect to σ are also easily normalized [Eberly 1994ab].

$$\hat{L}_{\sigma^k} = \hat{L}_{I,\sigma^k}(\mathbf{p} \mid \sigma) = \sigma^k \frac{\partial^k}{\partial \sigma^k} G(\sigma, \mathbf{p}) \otimes I(\mathbf{p}) \tag{3-5}$$

3.1.2. Reconstruction of Sampled Images via the Taylor Expansion

Combinations of derivative values can be used to recreate or approximate smooth functions from sparse samples. Digital images are discrete samplings of continuous functions. It is necessary to reconstruct a continuous function from the image samples in order to perform many image analysis operations. Given a sparse spatial sampling of the infinite set of derivatives of an image function (providing that the derivatives exist), it is possible to reconstruct a continuous function using a Taylor series expansion. For example, given a discretely sampled, continuously differentiable, 1D image function I(x), $x \in \mathsf{Z}$, a point $x_0 \in \mathsf{Z}$, and the set of all derivatives $D = \left\{\frac{\partial^n}{\partial x^n} I(x) \,\middle|\, n = 0, 1, 2, 3, \ldots\right\}$, the Taylor series can be used to reconstruct I(x), $x \in \mathsf{R}$, from *D*. The familiar series is shown in the following equation.

$\forall\, h \in \mathsf{R}$, and $x = x_0 + h$,

$$\begin{aligned} I(x) &= I(x_0 + h) \\ &= I(x_0) + h\left(\frac{\partial}{\partial x} I(x_0)\right) + \frac{h^2}{2!}\left(\frac{\partial^2}{\partial x^2} I(x_0)\right) + \cdots + \frac{h^k}{k!}\left(\frac{\partial^k}{\partial x^k} I(x_0)\right) + \cdots \\ &= \sum_{k=0}^{\infty} \frac{h^k}{k!}\left(\frac{\partial^k}{\partial x^k} I(x_0)\right) \end{aligned} \tag{3-6}$$

The instantaneous derivative values of an image function are seldom known; usually, only the zeroth order samples are available. Furthermore, the discrete sampling process produces an unrecoverable loss of information, governed by Shannon's sampling theorem and reflected by the Nyquist frequency. However, scaled derivatives of a discrete image can be measured, and a continuous representation of the image at some scale can be constructed. If I(x), $x \in \mathsf{Z}$ is the discrete 1D image function from the example above, let $L_{I,x^n}(x \mid \sigma)$ be the n-th order scale-space derivative of I(x) at scale σ, (i.e., the 1D analog

of the derivative values in equation (3-1)), and let the set of corresponding scale-space derivatives of I(x) be $L = \left\{ L_{I,x^n}(x \mid \sigma) \mid n = 0,1,2,3,\ldots \right\}$. The multiscale form of equation (3-6) is

$$\forall\ h \in R \text{ and } x = x_0 + h, \qquad L_I(x \mid \sigma) = \sum_{k=0}^{\infty} \frac{h^k}{k!} \left(L_{I,x^k}(x_0 \mid \sigma) \right) \tag{3-7}$$

Similarly, 2D scale-space images (i.e., $L_I(\mathbf{p} \mid \sigma)$ where $\mathbf{p} \in R^2$), can be reconstructed from multiscale derivatives.

$$\forall\ \mathbf{h} \in R^2,\ \mathbf{h} = (h_x\ h_y), \text{ and } \mathbf{p} = \mathbf{p}_0 + \mathbf{h}, \quad L_I(\mathbf{p} \mid \sigma) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{h_x{}^n h_y{}^m}{(n+m)!} \left(L_{I,x^n y^m}(\mathbf{p}_0 \mid \sigma) \right) \tag{3-8}$$

Eberly's scale-space 1-form requires that all spatial differences and scale differences be made relative to the scale at which the difference is measured. Therefore, the Taylor polynomials should be expressed in terms of scale σ, and they should use the normalized dimensionless derivative values $\hat{L}_{x^p y^q}$. For a 1D image function, let $\hat{h} = h/\sigma$ (Thus, $\hat{h}\sigma = h$). Transforming the previous Taylor polynomial in equation (3-7) to the dimensionless offset value of $\hat{h}$ yields

$$L_I(x \mid \sigma) = \sum_{k=0}^{\infty} \frac{\hat{h}^k}{k!} \left(\hat{L}_{I,x^k}(x_0 \mid \sigma) \right) \tag{3-9}$$

The 2D offset $\mathbf{h} = (h_x, h_y)$ can be normalized to $\hat{\mathbf{h}} = (\hat{h}_x, \hat{h}_y) = (\frac{h_x}{\sigma}, \frac{h_y}{\sigma})$. The result is a transformation of equation (3-8) to a corresponding dimensionless Taylor expression.

$$L_I(\mathbf{p} \mid \sigma) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{\hat{h}_x{}^n \hat{h}_y{}^m}{n!m!} \left(\hat{L}_{I,x^n y^m}(\mathbf{x}_0 \mid \sigma) \right) \tag{3-10}$$

3.1.3. Exploring the Properties of Scale-Space Derivatives

In essence, this chapter is about understanding scale-space derivatives, their uses, and their properties. The preceding section has presented scale-space derivatives in both unnormalized and normalized forms, and supplied some insight into their uses. Scale-space derivatives provide a vehicle for reconstructing smooth image functions at some scale.

The next step in the exploration of scale-space derivatives is the understanding their noise properties. What are the relations between one derivative and another? How do these relationships change as scale increases?

More precisely, derivatives are compared according to the way that they propagate noise from the original image through changing scale. How sensitive is a multiscale derivative to spatially uncorrelated white noise in the original digital image signal? How does this sensitivity compare with the response of other derivatives of different order?

This chapter explores both scale-space differential forms $L_{x^p y^q}$ and $\hat{L}_{x^p y^q}$ and their interactions with noisy input across scale.

3.2. Noise and Scale

Consider a 1D image with added white noise; that is, let $\tilde{I}(x) = I(x) + \tilde{u}$ such that $x \in \mathbb{R}$, and $\tilde{u}$ is a zero-mean, spatially uncorrelated normally distributed random variable with variance v_0 (i.e., the probability distribution function of $\tilde{u} = N_{0,v_0}(\tilde{u})$). As a linear function of a random variable $\tilde{u}$, $\tilde{I}(x)$ can be expressed as function $I(x,\tilde{u})$ whose mean $M(I(x,\tilde{u})) = M(\tilde{I}(x))$ (or the first moment of $\tilde{I}(x)$, $\mu_{\tilde{I}}(x)$) and variance $V(I(x,\tilde{u})) = V(\tilde{I}(x))$ (or the second central moment $\mu_{\tilde{I}}^{(2)}(x)$ of $\tilde{I}(x)$,) are calculated as shown in Chapter 2, equations (2-12) and (2-13). The following two equations revisit those earlier calculations.

$$\begin{aligned} M(I(x,\tilde{u})) = M(\tilde{I}(x)) &= \mu_{\tilde{I}}(x) = \langle I(x,\tilde{u}) \rangle \\ &= \int (I(x) + \tilde{u})(N_{0,v_0}(\tilde{u}))d\tilde{u} \\ &= I(x)\int (N_{0,v_0}(\tilde{u}))d\tilde{u} + \langle \tilde{u} \rangle \\ &= I(x) \end{aligned} \tag{3-11}$$

$$\begin{aligned} V(I(x,\tilde{u})) = V(\tilde{I}(x)) &= \mu_{\tilde{I}}^{(2)}(x) \\ &= \langle (I(x,\tilde{u}) - I(x))^2 \rangle \\ &= \int \tilde{u}(N_{0,v_0}(\tilde{u}))d\tilde{u} \\ &= v_0 \end{aligned} \tag{3-12}$$

Applying a multiscale evaluation of the image $\tilde{I}(x)$ extends the forms of $\mu_{\tilde{I}}(x)$ and $\mu_{\tilde{I}}^{(2)}(x)$ to include a scale parameter. Consider the convolution of $I(x,\tilde{u})$ with an arbitrary filter kernel h(x).

$$\begin{aligned} \mu_{\tilde{I}\otimes h}^{(2)}(x) &= V\left(\int_{-\infty}^{\infty} I(x-\tau,\tilde{u})h(\tau)d\tau\right) \\ &= \int_{-\infty}^{\infty} (h(\tau))^2 V(I(x-\tau,\tilde{u}))d\tau \\ &= v_0 \int_{-\infty}^{\infty} (h(\tau))^2 d\tau \end{aligned} \tag{3-13}$$

The variance of $I(x,\tilde{u})$, convolved with a filter kernel h(x) is dependent on the structure of the kernel, and not on the underlying function I(x). This relation is true for all functions $I(x,\tilde{u})$ with zero-mean Gaussian additive spatially uncorrelated white noise and all filter kernels h(x) for any n-dimensional space.

Combining the zeroth order scale-space derivative $L_{\tilde{I}}(x \mid \sigma)$ and equation (3-13) results in the following relations for 1D images.

$$\begin{aligned} V\left(L_{\tilde{I}}(x \mid \sigma)\right) &= \mu^{(2)}_{\tilde{I}}(x \mid \sigma) = V\left(\int_{-\infty}^{\infty} \tilde{I}(x-\tau)G(\sigma, x)d\tau\right) \\ &= v_0 \int_{-\infty}^{\infty} (G(\sigma, x))^2 d\tau \\ &= v_0 \int_{-\infty}^{\infty} \left(\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} \right)^2 d\tau \\ &= v_0 \frac{1}{2\sigma\sqrt{\pi}} \int_{-\infty}^{\infty} \frac{1}{(\sigma/\sqrt{2})\sqrt{2\pi}} e^{-\frac{x^2}{2(\sigma/\sqrt{2})^2}} d\tau \\ &= \frac{v_0}{2\sigma\sqrt{\pi}} \end{aligned} \tag{3-14}$$

$$\begin{aligned} M\left(L_{\tilde{I}}(x \mid \sigma)\right) &= \mu_{\tilde{I}}(x \mid \sigma) = \langle I(x, \tilde{u}) \otimes G(\sigma, x) \rangle \\ &= \int\int_{-\infty}^{\infty} (I(x-\tau) + \tilde{u}) G(\sigma, x) d\tau \left(N_{0, v_0}(\tilde{u})\right) d\tilde{u} \\ &= (I(x) \otimes G(\sigma, x)) \int \left(N_{0, v_0}(\tilde{u})\right) d\tilde{u} + \langle \tilde{u} \rangle \\ &= I(x) \otimes G(\sigma, x) \end{aligned} \tag{3-15}$$

Thus the variance of the scale-space zeroth order intensity values is inversely proportional to scale σ and directly proportional to the variance of the noise in the input image. Keep in mind that the variance so described is distributed about the mean value of the scale-space intensity measurement and is thus a measure of the error relative to the scale at which it is measured and centered about the expected intensity value.

The relation described in (3-14) generalizes to higher dimensions. If the initial image is 2D and the Gaussian scale-space sampling kernel is also 2D, the relation in (3-14) can be rewritten as

$$V\left(L_{\tilde{I}}(\mathbf{p} \mid \sigma)\right) = \mu^{(2)}_{\tilde{I}}(\mathbf{p} \mid \sigma) = \frac{v_0}{4\pi\sigma^2}, \ \mathbf{p} \in R^2 \tag{3-16}$$

3.3. Variance of Multiscale Derivatives without Normalization

In his dissertation research, Blom performed a statistical analysis on scale-space derivatives of arbitrary order. He measured the response of the multiscale derivative measurement to noise as a function of scale [Blom 1992]. Blom summarizes his results using 2D images as the basis of his analysis. For the purposes of this discussion, Blom's analysis is recreated here for images of one dimension by extending the zero-th order statistical relations in equations (3-14) and (3-15) to unnormalized 1D scale-space derivatives.

3.3.1. Covariances of 1D Multiscale Derivatives

This is a construction of multiscale covariances for Gaussian scale space of 1D images with Gaussian additive white noise. These results are not original observations. Similar findings are reported by Metz [Metz 1969] and by Blom [Blom 1992] as well as elsewhere in the literature.

Given the previously defined 1D image function with additive Gaussian white noise $\tilde{I}(x)$, the covariance of two scale-space derivatives of $\tilde{I}(x)$ can be measured for any location x at scale σ. The covariance between two such derivatives is defined as

$$\begin{aligned}&\mathrm{Cov}\left(L_{\tilde{I},x^i}(x\mid\sigma),\, L_{\tilde{I},x^j}(x\mid\sigma)\right)\\&\quad=\left\langle\left(L_{\tilde{I},x^i}(x\mid\sigma)-M(L_{\tilde{I},x^i}(x\mid\sigma))\right)\left(L_{\tilde{I},x^j}(x\mid\sigma)-M(L_{\tilde{I},x^j}(x\mid\sigma))\right)\right\rangle\end{aligned}\qquad(3\text{-}17)$$

where angle brackets indicated the expected value operation and $M(L_{\tilde{I},x^i}(x\mid\sigma))$ is the mean or expected value of the i-th scale-space derivative of $\tilde{I}(x)$. Observing that

$$\begin{aligned}M(L_{\tilde{I},x^i}(x\mid\sigma))&=\left\langle\tfrac{\partial^i}{\partial x^i}G(\sigma,x)\otimes\left(I(x)+\tilde{u}\right)\right\rangle\\&=\left\langle\tfrac{\partial^i}{\partial x^i}G(\sigma,x)\otimes I(x)\right\rangle+\left\langle\tfrac{\partial^i}{\partial x^i}G(\sigma,x)\otimes\tilde{u}\right\rangle\\&=\tfrac{\partial^i}{\partial x^i}G(\sigma,x)\otimes I(x)\\&=L_{I,x^i}(x\mid\sigma)\end{aligned}\qquad(3\text{-}18)$$

and recalling that convolution distributes over addition yields the following simplification:

$$\begin{aligned}&\mathrm{Cov}\left(L_{\tilde{I},x^i}(x\mid\sigma)\,,\,L_{\tilde{I},x^j}\tilde{I}(x\mid\sigma)\right)\\&\quad=\left\langle\left(L_{\tilde{I},x^i}(x\mid\sigma)-L_{I,x^i}(x\mid\sigma)\right)\left(L_{\tilde{I},x^j}\tilde{I}(x\mid\sigma)-L_{I,x^j}\tilde{I}(x\mid\sigma)\right)\right\rangle\\&\quad=\left\langle\left(\tfrac{\partial^i}{\partial x^i}G(\sigma,x)\otimes(I(x)+\tilde{u})-\tfrac{\partial^i}{\partial x^i}G(\sigma,x)\otimes I(x)\right)\right.\\&\qquad\left.\left(\tfrac{\partial^j}{\partial x^j}G(\sigma,x)\otimes(I(x)+\tilde{u})-\tfrac{\partial^j}{\partial x^j}G(\sigma,x)\otimes I(x)\right)\right\rangle\\&\quad=\left\langle\left(\tfrac{\partial^i}{\partial x^i}G(\sigma,x)\otimes\tilde{u}\right)\left(\tfrac{\partial^j}{\partial x^j}G(\sigma,x)\otimes\tilde{u}\right)\right\rangle\end{aligned}\qquad(3\text{-}19)$$

Replacing the convolution operator with its corresponding integral and simplifying yields

$$
\begin{aligned}
&\mathrm{Cov}\left(L_{\tilde{I},x^i}(x\,|\,\sigma),\, L_{\tilde{I},x^j}(x\,|\,\sigma)\right) \\
&\quad = \left\langle \int_{-\infty}^{\infty} \tfrac{\partial^i}{\partial x^i} G(\sigma,\tau)\tilde{u}(x-\tau)\,d\tau \int_{-\infty}^{\infty} \tfrac{\partial^j}{\partial x^j} G(\sigma,\nu)\tilde{u}(x-\nu)\,d\nu \right\rangle \\
&\quad = \left\langle \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \tfrac{\partial^i}{\partial x^i} G(\sigma,\tau) \tfrac{\partial^j}{\partial x^j} G(\sigma,\nu)\tilde{u}(x-\tau)\tilde{u}(x-\nu)\,d\nu\,d\tau \right\rangle \\
&\quad = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \tfrac{\partial^i}{\partial x^i} G(\sigma,\tau) \tfrac{\partial^j}{\partial x^j} G(\sigma,\nu)\left\langle \tilde{u}(x-\tau)\tilde{u}(x-\nu)\right\rangle d\nu\,d\tau
\end{aligned}
\tag{3-20}
$$

The term $\langle \tilde{u}(x-\tau)\tilde{u}(x-\nu)\rangle$ is by definition the spatial correlation of the additive noise function $\tilde{u}$ relative to the location x. Since $\tilde{u}$ is assumed to be white, its distribution is Gaussian about a zero mean, and it is not correlated in space. That is

$$
\left\langle \tilde{u}(x-\tau)\tilde{u}(x-\nu)\right\rangle = \begin{cases} \left\langle \left(\tilde{u}(x-\tau)\right)^2 \right\rangle = v_0 & , \quad \text{if } \nu = \tau \\ 0 & , \quad \text{otherwise} \end{cases}
\tag{3-21}
$$

Applying the results from (3-21), the covariance relation reduces to

$$
\begin{aligned}
\mathrm{Cov}\left(L_{\tilde{I},x^i}(x\,|\,\sigma),\, L_{\tilde{I},x^j}(x\,|\,\sigma)\right) &= \int_{-\infty}^{\infty} \tfrac{\partial^i}{\partial x^i} G(\sigma,\tau) \tfrac{\partial^j}{\partial x^j} G(\sigma,\tau)\left\langle \left(\tilde{u}(x-\tau)\right)^2 \right\rangle d\tau \\
&= \int_{-\infty}^{\infty} \tfrac{\partial^i}{\partial x^i} G(\sigma,\tau) \tfrac{\partial^j}{\partial x^j} G(\sigma,\tau) v_0\,d\tau \\
&= v_0 \int_{-\infty}^{\infty} \tfrac{\partial^i}{\partial x^i} G(\sigma,\tau) \tfrac{\partial^j}{\partial x^j} G(\sigma,\tau)\,d\tau
\end{aligned}
\tag{3-22}
$$

Simplifying the integral in equation (3-22) requires the use of several identities involving Hermite polynomials. A complete derivation is provided in the appendix of this chapter. The resulting simplified relation is

$$
\begin{aligned}
&\mathrm{Cov}\left(L_{\tilde{I},x^i}(x\,|\,\sigma),\, L_{\tilde{I},x^j}(x\,|\,\sigma)\right) \\
&= \begin{cases} (-1)^{i+\left(\frac{i+j}{2}\right)} \dfrac{v_0}{2\sigma\sqrt{\pi}} \left(\dfrac{1}{\sigma\sqrt{2}}\right)^{i+j} \left(\displaystyle\prod_{r=0}^{(i+j)/2} (2r-1) \right) & , \forall \text{ even } i+j \\ 0 & , \forall \text{ odd } i+j \end{cases}
\end{aligned}
\tag{3-23}
$$

Variance is a special case of covariance. Using equation (3-23), the general form for the variance of any k-th order (k > 0) unnormalized scale-space derivative of a 1D image is shown to be

$$
V\left(L_{\tilde{I},k}(x\,|\,\sigma)\right) = \frac{v_0}{2\sigma\sqrt{\pi}} \frac{\displaystyle\prod_{i=1}^{k}(2i-1)}{2^k \sigma^{2k}}
\tag{3-24}
$$

Values for the variances of scale-space derivatives (order 0 - 6) are shown in Table 3.1.

Derivative	Variance
$L_{\tilde{I}}(x \mid \sigma)$	$\frac{v_0}{2\sigma\sqrt{\pi}}$
$L_{\tilde{I},x}(x \mid \sigma)$	$\frac{1}{2\sigma^2}\frac{v_0}{2\sigma\sqrt{\pi}}$
$L_{\tilde{I},x^2}(x \mid \sigma)$	$\frac{3}{4\sigma^4}\frac{v_0}{2\sigma\sqrt{\pi}}$
$L_{\tilde{I},x^3}(x \mid \sigma)$	$\frac{15}{8\sigma^6}\frac{v_0}{2\sigma\sqrt{\pi}}$
$L_{\tilde{I},x^4}(x \mid \sigma)$	$\frac{105}{16\sigma^8}\frac{v_0}{2\sigma\sqrt{\pi}}$
$L_{\tilde{I},x^5}(x \mid \sigma)$	$\frac{945}{32\sigma^{10}}\frac{v_0}{2\sigma\sqrt{\pi}}$
$L_{\tilde{I},x^6}(x \mid \sigma)$	$\frac{10395}{64\sigma^{12}}\frac{v_0}{2\sigma\sqrt{\pi}}$

Table 3.1. Variances of unnormalized scale-space derivatives (order 0-6) of noisy 1D images (variance of input noise = v_0)

3.3.2. Covariances of 2D Multiscale Derivatives

Blom and ter Haar Romeny [Blom 1992][ter Haar Romeny 1993] present similar results for unnormalized scale-space derivatives of noisy 2D images (Gaussian distributed, zero mean additive noise). Their results are summarized in Table 3.2.

Derivative	Variance
$L_{\tilde{I}}(\mathbf{p} \mid \sigma)$	$\frac{v_0}{4\sigma^2\pi}$
$L_{\tilde{I},x}(\mathbf{p} \mid \sigma), L_{\tilde{I},y}(\mathbf{p} \mid \sigma)$	$\frac{1}{4\sigma^2}\frac{v_0}{4\sigma^2\pi}$
$L_{\tilde{I},x^2}(\mathbf{p} \mid \sigma), L_{\tilde{I},y^2}(\mathbf{p} \mid \sigma)$	$\frac{3}{16\sigma^4}\frac{v_0}{4\sigma^2\pi}$
$L_{\tilde{I},xy}(\mathbf{p} \mid \sigma)$	$\frac{1}{16\sigma^4}\frac{v_0}{4\sigma^2\pi}$
$L_{\tilde{I},x^3}(\mathbf{p} \mid \sigma), L_{\tilde{I},y^3}(\mathbf{p} \mid \sigma)$	$\frac{15}{64\sigma^6}\frac{v_0}{4\sigma^2\pi}$
$L_{\tilde{I},x^2y}(\mathbf{p} \mid \sigma), L_{\tilde{I},xy^2}(\mathbf{p} \mid \sigma)$	$\frac{3}{64\sigma^6}\frac{v_0}{4\sigma^2\pi}$
$L_{\tilde{I},x^4}(\mathbf{p} \mid \sigma), L_{\tilde{I},y^4}(\mathbf{p} \mid \sigma)$	$\frac{105}{256\sigma^8}\frac{v_0}{4\sigma^2\pi}$
$L_{\tilde{I},x^3y}(\mathbf{p} \mid \sigma), L_{\tilde{I},xy^3}(\mathbf{p} \mid \sigma)$	$\frac{15}{256\sigma^8}\frac{v_0}{4\sigma^2\pi}$
$L_{\tilde{I},x^2y^2}(\mathbf{p} \mid \sigma)$	$\frac{9}{256\sigma^8}\frac{v_0}{4\sigma^2\pi}$

Table 3.2. Variances of unnormalized scale-space derivatives of noisy 2D images (variance of input noise = v_0) for partial spatial derivatives to the fourth order (Adapted from Blom 1992)

The general form for the variance of unnormalized scale-space derivatives of a two-dimensional image, where j represents the order of differentiation in the x direction and k is the order of differentiation in the y direction (j > 0; k > 0) is given by

$$V\left(L_{\hat{I},x^j y^k}(\mathbf{p}\mid\sigma)\right)=\frac{v_0}{4\sigma^2\pi}\frac{\prod_{n=1}^{i}(2n-1)\prod_{m=1}^{j}(2m-1)}{2^{(j+k)}\sigma^{2(j+k)}} \tag{3-25}$$

3.4. Variance of Normalized Scale-Space Derivatives

The values shown in Tables 3.1 and 3.2 are a reflection of the absolute propagated error present in the resulting scale-space derivative images at a single scale σ. However, if a measurement is to be made *across* different scales (i.e., comparing derivative results at two different sampling/filter apertures) the derivative values must be normalized to ensure measurements that are invariant with respect to changing scale. Subsequently, the statistics of these measurements, when made of noisy images, must also reflect the normalization. This section describes new observations of normalized scale-space differential invariants.

Using the relations found in (3-3) and (3-24), it is straightforward to determine the variance of normalized 1D scale-space derivatives $V(\hat{L}_{\hat{I},k}(x\mid\sigma))$.

$$V(\hat{L}_{\hat{I},k}(x\mid\sigma))=V(\sigma^k L_{\hat{I},k}(x\mid\sigma))=\sigma^{2k}V(L_{\hat{I},k}(x\mid\sigma)) \tag{3-26}$$

Substituting from (3-24) into equation (3-26) yields the following general form for the variance of normalized derivatives of 1D images:

$$V\left(\hat{L}_{\hat{I},k}(x\mid\sigma)\right)=\frac{v_0}{2\sigma\sqrt{\pi}}\frac{\prod_{i=1}^{k}(2i-1)}{2^k} \tag{3-27}$$

Using this relation, the results of Table 3.1 are recalculated for normalized spatial derivatives and shown in Table 3.3.

Similarly, the general form for the variance of normalized scale-space derivatives of a two-dimensional image is given by

$$V\left(\hat{L}_{\hat{I},x^j y^k}(\mathbf{p}\mid\sigma)\right)=\frac{v_0}{4\sigma^2\pi}\frac{\prod_{n=1}^{i}(2n-1)\prod_{m=1}^{j}(2m-1)}{2^{(j+k)}} \tag{3-28}$$

where **p** =(x y), j represents the order of differentiation in the x direction, and k is the order of differentiation in the y direction (j > 0; k > 0). The resulting variance statistics are generated for normalized scale-space derivatives of 2D images up to the fourth order are presented in Table 3.4.

Derivative	Variance
$\hat{L}_{\tilde{I}}(x \mid \sigma)$	$\frac{v_0}{2\sigma\sqrt{\pi}}$
$\hat{L}_{\tilde{I},x}(x \mid \sigma)$	$\frac{1}{2}\frac{v_0}{2\sigma\sqrt{\pi}}$
$\hat{L}_{\tilde{I},x^2}(x \mid \sigma)$	$\frac{3}{4}\frac{v_0}{2\sigma\sqrt{\pi}}$
$\hat{L}_{\tilde{I},x^3}(x \mid \sigma)$	$\frac{15}{8}\frac{v_0}{2\sigma\sqrt{\pi}}$
$\hat{L}_{\tilde{I},x^4}(x \mid \sigma)$	$\frac{105}{16}\frac{v_0}{2\sigma\sqrt{\pi}}$
$\hat{L}_{\tilde{I},x^5}(x \mid \sigma)$	$\frac{945}{32}\frac{v_0}{2\sigma\sqrt{\pi}}$
$\hat{L}_{\tilde{I},x^6}(x \mid \sigma)$	$\frac{10395}{64}\frac{v_0}{2\sigma\sqrt{\pi}}$

Table 3.3. Variances of normalized scale-space derivatives (order 0-6) of noisy 1D images (variance of input noise = v_0)

Derivative	Variance
$\hat{L}_{\tilde{I}}(\mathbf{p} \mid \sigma)$	$\frac{v_0}{4\sigma^2\pi}$
$\hat{L}_{\tilde{I},x}(\mathbf{p} \mid \sigma), \hat{L}_{\tilde{I},y}(\mathbf{p} \mid \sigma)$	$\frac{1}{4}\frac{v_0}{4\sigma^2\pi}$
$\hat{L}_{\tilde{I},x^2}(\mathbf{p} \mid \sigma), \hat{L}_{\tilde{I},y^2}(\mathbf{p} \mid \sigma)$	$\frac{3}{16}\frac{v_0}{4\sigma^2\pi}$
$\hat{L}_{\tilde{I},xy}(\mathbf{p} \mid \sigma)$	$\frac{1}{16}\frac{v_0}{4\sigma^2\pi}$
$\hat{L}_{\tilde{I},x^3}(\mathbf{p} \mid \sigma), \hat{L}_{\tilde{I},y^3}(\mathbf{p} \mid \sigma)$	$\frac{15}{64}\frac{v_0}{4\sigma^2\pi}$
$\hat{L}_{\tilde{I},x^2y}(\mathbf{p} \mid \sigma), \hat{L}_{\tilde{I},xy^2}(\mathbf{p} \mid \sigma)$	$\frac{3}{64}\frac{v_0}{4\sigma^2\pi}$
$\hat{L}_{\tilde{I},x^4}(\mathbf{p} \mid \sigma), \hat{L}_{\tilde{I},y^4}(\mathbf{p} \mid \sigma)$	$\frac{105}{256}\frac{v_0}{4\sigma^2\pi}$
$\hat{L}_{\tilde{I},x^3y}(\mathbf{p} \mid \sigma), \hat{L}_{\tilde{I},xy^3}(\mathbf{p} \mid \sigma)$	$\frac{15}{256}\frac{v_0}{4\sigma^2\pi}$
$\hat{L}_{\tilde{I},x^2y^2}(\mathbf{p} \mid \sigma)$	$\frac{9}{256}\frac{v_0}{4\sigma^2\pi}$

Table 3.4. Variances of normalized scale-space derivatives of noisy 2D images for partial spatial derivatives to the fourth order (variance of input noise = v_0)

3.5. Analysis of 1D Scale-Space Derivatives

While the structure of equations (3-24) and (3-27) are very similar (indeed, the coefficients are identical), the variance of normalized scale-space derivatives are inversely proportional to σ, while the unnormalized derivatives are inversely proportional to σ^{2k+1} (k is the order of differentiation). This difference can be shown graphically by

plotting the propagated error of each of the derivative measurements vs. the scale parameter σ for different orders of differentiation.

Figure 3.1 shows the propagated error of scale-space derivatives without normalization. Figure 3.2 shows the propagated error with the normalizing factor included. In both cases, the plots are on a Log-Log scale. The order of differentiation k shown by the different curves ranges from k = 0 to k = 6 and is labeled on the far left of each figure.

Comparing the two plots, the most important difference is immediately clear. In both plots, the propagated error in either normalized or unnormalized scale-space derivatives is consistently decreasing with increasing scale σ. However, in the unnormalized measurements there is a significant crossing where for relatively large scale ($\sigma > 1$), the variance (and subsequently the significance) of the propagated error is smaller for higher order derivatives.

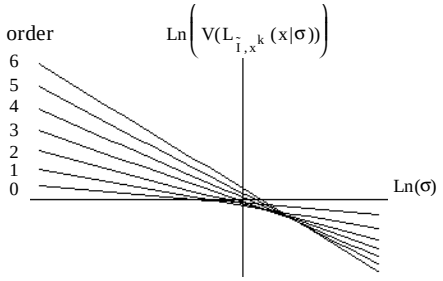

Figure 3.1. Propagated error of unnormalized 1D scale-space derivatives (order 0-6). Each curve represents the ratio of variance of output to input noise of the linear unnormalized derivative of Gaussian filter vs. scale σ. Plot is on a log-log scale.

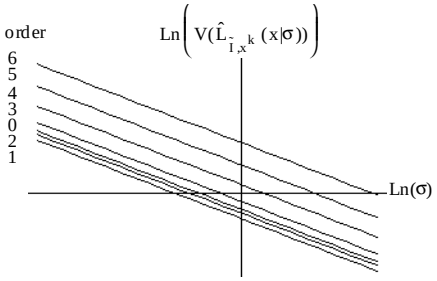

Figure 3.2 Plot of the propagated error of normalized 1D scale-space derivatives (order 0-6). Each curve represents the ratio of variance of output to input noise of the linear normalized derivative of Gaussian filter vs. scale σ. Plot is on a log-log scale.

The crossover exhibited in the unnormalized derivatives has been used to justify the application of very high order derivative filters in the analysis and reconstruction of images. When comparing across scale, a normalized representation is required. If the derivatives are normalized, the improved stability of derivative measurements as scale increases remains, but the assertion that the relative stability and accuracy improves with increasing order of differentiation at large scale does not hold.

An alternate finding for normalized scale-space derivatives becomes apparent. For 1D normalized scale-space derivatives, the low order differential forms (k=1 and k=2) propagate less noise than either the luminance (zeroth order form or k=0) or any of the derivatives for k > 2. Figure 3.3 shows the noise propagation of normalized scale-space derivatives versus order of differentiation relative to the zeroth order or scale-space luminance noise. The 'J' shape of the curve bears consideration (see Section 3.7, "Discussion").

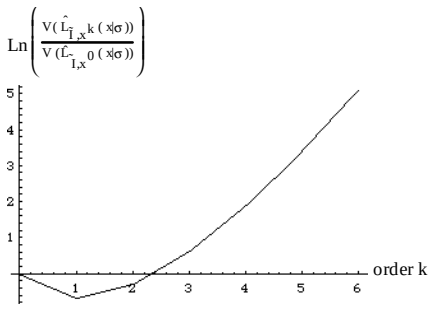

Figure 3.3. Plot of the propagated error of normalized 1D scale-space derivatives (order 0-6). Curve represents the ratio of variance of output to input noise of the linear unnormalized derivative of Gaussian filter vs. order of differentiation. Plot is on a log scale.

3.6. Analysis of 2D Scale-Space Derivatives

To create a comparable analysis for 2D images, the appropriate variances must combine the contributions of all the partial derivatives of the same order. Consider the 2D Taylor expansion for a scale-space function $L_{\tilde{I}}(\mathbf{p} \mid \sigma)$. Let $\mathbf{p}_0 \in \mathsf{R}^2$, $\mathbf{h} \in \mathsf{R}^2$ such that the interval $\mathbf{h} = (h_x, h_y)$, and $\mathbf{p} \in \mathsf{R}^2$ where $\mathbf{p} = \mathbf{p}_0 + \mathbf{h}$. Rewriting equation (3-8) as an expression of the order of differentiation k generates the following representation.

$$\begin{aligned} L_{\tilde{I}}(\mathbf{p} \mid \sigma) = L_{\tilde{I}}(\mathbf{p}_0 \mid \sigma) + h_x L_{I,x}(\mathbf{p}_0 \mid \sigma) + h_y L_{I,y}(\mathbf{p}_0 \mid \sigma) + \cdots \\ + \frac{1}{k!} \sum_{i=0}^{k} h_x{}^i h_y{}^{k-i} \left(L_{I,x^i y^{k-i}}(\mathbf{p}_0 \mid \sigma) \right) + \cdots \end{aligned} \tag{3-29}$$

Inspection the k-th order term shows that each partial derivative contributes to the k-th value weighted by the interpolation interval. The 2D variance analog to the 1D treatment shown before requires the measurement of the variance of the k-th order term. These calculations require the combination of covariances of partial derivatives of the same order. Blom derived an expression for the covariance of two partial derivatives of 2D images [Blom 1992]. Formally the expression of the variance of the k-th order term of the 2D scale-space Taylor expansion is shown below in equation (3-30).

$$\begin{aligned} V(T(k)) &= V\left(\sum_{n=0}^{k} h_x{}^n h_y{}^{k-n} L_{\tilde{I},x^n y^{k-n}}(\mathbf{p}_0 \mid \sigma) \right) \\ &= \sum_{i=0}^{k} \sum_{j=0}^{k} h_x{}^{k+i-j} h_y{}^{k+j-i} \mathrm{Cov}\left(L_{\tilde{I},x^i y^{k-i}}(\mathbf{p}_0 \mid \sigma), L_{\tilde{I},x^{k-j} y^j}(\mathbf{p}_0 \mid \sigma) \right) \\ &= \frac{v_0}{4\sigma^2 \pi} \frac{1}{2^k \sigma^{2k}} \sum_{i=0}^{k} \sum_{j=0}^{k} h_x{}^{k+i-j} h_y{}^{k+j-i} \prod_{n=1}^{(k+i-j)/2} (2n-1) \prod_{m=1}^{(k-i+j)/2} (2m-1) \\ &\qquad \forall \text{ even } k + i - j \text{ and even } k - i + j \end{aligned} \tag{3-30}$$

To set the offset vector **h** to a common value, consider all values equidistant from the point $\mathbf{p}_0$. That is, consider for some radius r the set of all points $\{\mathbf{h} = (h_x, h_y) \mid r^2=h_x{}^2+h_y{}^2\}$. Therefore, let $h_x = r\cos(\theta)$ and $h_y = r\sin(\theta)$. Substitute these values into equation (3-30).

$$V\left(\sum_{n=0}^{k}(r\cos\theta)^n(r\sin\theta)^{k-n}L_{\tilde{I},x^n y^{k-n}}(\mathbf{p}_0 \mid \sigma)\right)$$
$$= \frac{v_0}{4\sigma^2\pi}\frac{r^{2k}}{2^k\sigma^{2k}}\sum_{i=0}^{k}\sum_{j=0}^{k}\cos(\theta)^{k+i-j}\sin(\theta)^{k+j-i}\prod_{n=1}^{(k+i-j)/2}(2n-1)\prod_{m=1}^{(k-i+j)/2}(2m-1)$$
$$\forall \text{ even } k+i-j \text{ and even } k-i+j \qquad (3\text{-}31)$$

Evaluating the variance expression in terms of θ shows that such a representation of variance is cyclic for all values of r and all values of σ. The maximum value of the variance expression is found when $\theta = 0$, $\theta = \frac{1}{2}\pi$, $\theta = \pi$, $\theta = \frac{3}{2}\pi$, or $\theta = 2\pi$. This implies that without a loss of generality, the error propagated from uncorrelated noise through unnormalized 2D scaled derivatives can be bounded by

$$\mathrm{Max}_\theta\left(V\left(\sum_{n=0}^{k}(r\cos\theta)^n(r\sin\theta)^{k-n}L_{\tilde{I},x^n y^{k-n}}(\mathbf{p}_0 \mid \sigma)\right)\right) = \frac{v_0}{4\sigma^2\pi}\frac{r^{2k}}{2^k\sigma^{2k}}\prod_{n=1}^{k}(2n-1) \qquad (3\text{-}32)$$

Given $h_x = r\cos(\theta)$ and $h_y = r\sin(\theta)$, consider the variance of the combined k-th order terms of the 2D Taylor expansion relative to the magnitude of **h**. That is, consider

$$V_{max}(T(k)) = \left(\frac{1}{r^{2k}}\right)\mathrm{Max}_\theta\left(V\left(\sum_{n=0}^{k}(r\cos\theta)^n(r\sin\theta)^{k-n}L_{\tilde{I},x^n y^{k-n}}(\mathbf{p} \mid \sigma)\right)\right)$$
$$= \frac{v_0}{4\sigma^2\pi}\frac{1}{2^k\sigma^{2k}}\prod_{n=1}^{k}(2n-1) \qquad (3\text{-}33)$$

where T(k) represents the contribution of the aggregate k-th order derivative terms of the Taylor expansion, divided by the weighting of the magnitude of the interpolation offset **h**. The value in equation (3-33) represents the 2D analog of the variance of 1D scaled derivatives shown in equation (3-24).

If the 2D offset vector **h** is normalized to $\hat{\mathbf{h}} = (\hat{h}_x, \hat{h}_y) = (\frac{h_x}{\sigma}, \frac{h_y}{\sigma})$, then it follows that given $\hat{r} = r/\sigma$, $\hat{h}_x = \hat{r}\cos(\theta)$ and $\hat{h}_y = \hat{r}\sin(\theta)$. Repeating the analysis shown above for the dimensionless Taylor expansion shown in equation (3-10) and using the scaled quantities for $\hat{\mathbf{h}}$ and $\hat{r}$ generates an expression for the upper bound on propagated error from uncorrelated noise through normalized scaled derivatives:

$$\mathrm{Max}_\theta\left(V\left(\sum_{n=0}^{k}(r\cos\theta)^n(r\sin\theta)^{k-n}\hat{L}_{\tilde{I},x^n y^{k-n}}(\mathbf{p} \mid \sigma)\right)\right) = \frac{v_0}{4\sigma^2\pi}\frac{\hat{r}^{2k}}{2^k}\prod_{n=1}^{k}(2n-1) \qquad (3\text{-}34)$$

If $\hat{T}(k)$ is the aggregate contribution of the k-th order normalized derivatives to the dimensionless Taylor expansion to be weighted by the normalized magnitude $\hat{r} = r/\sigma$, then the normalized 2D variances of the k-th order, unweighted by $\hat{r}$ is

$$V_{max}\left(\hat{T}(k)\right) = \left(\frac{1}{\hat{r}^{2k}}\right) Max\left(V\left(\sum_{n=0}^{k} (\hat{r}\cos\theta)^n (\hat{r}\sin\theta)^{k-n} \hat{L}_{I, x^n y^{k-n}}(\mathbf{p} \mid \sigma) \right) \right)$$
$$= \frac{v_0}{4\sigma^2\pi} \frac{1}{2^k} \prod_{n=1}^{k} (2n-1) \qquad (3\text{-}35)$$

The results for both the normalized and unnormalized expressions are summarized in Table 3.5.

Order (k)	$V_{max}(T(k))$	$V_{max}\left(\hat{T}(k)\right)$
0	$\frac{v_0}{4\pi\sigma^2}$	$\frac{v_0}{4\pi\sigma^2}$
1	$\frac{v_0}{8\pi\sigma^4}$	$\frac{v_0}{8\pi\sigma^2}$
2	$\frac{3v_0}{16\pi\sigma^6}$	$\frac{3v_0}{16\pi\sigma^2}$
3	$\frac{15v_0}{32\pi\sigma^8}$	$\frac{15v_0}{32\pi\sigma^2}$
4	$\frac{105v_0}{64\pi\sigma^{10}}$	$\frac{105v_0}{64\pi\sigma^2}$
5	$\frac{945v_0}{128\pi\sigma^{12}}$	$\frac{945v_0}{128\pi\sigma^2}$
6	$\frac{10395v_0}{256\pi\sigma^{14}}$	$\frac{10395v_0}{256\pi\sigma^2}$

Table 3.5. Variances of both unnormalized and normalized scale-space derivatives (order 0-6) of noisy 2D images (variance of input noise = v_0)

The equations of Table 3.5 are easily plotted and their results portrayed graphically in Figures 3.4 and 3.5. These plots resemble their 1D counterparts, and the conclusions drawn from them are the same as in the 1D case. Recapitulating the 1D results, in all cases the variances are monotonically decreasing with scale; however, if normalized values are considered in order to make cross-scale comparisons, the crossing of the variance values of the derivative values through increasing scale does not persist.

As in the 1D case, under normalization the values of the variances of 2D derivatives are not strictly increasing with rising order of differentiation k. The first and second order normalized 2D derivatives propagate less noise than the scaled zeroth order intensities, regardless of scale. Figure 3.6 shows variances of 2D derivative terms $\hat{T}(k)$ plotted as a function of the order of differentiation and relative to the variance of the zeroth order scaled intensity value $\hat{T}(0)$.

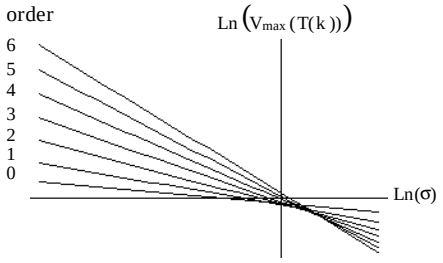

Figure 3.4. Propagated error of unnormalized 2D scale-space derivatives (order 0-6). Each curve represents the ratio of variance of output to input noise of the linear unnormalized derivative of Gaussian filter vs. scale σ. Plot is on a log-log scale.

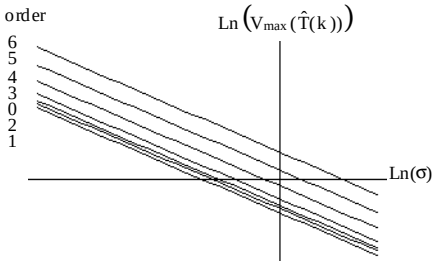

Figure 3.5 Plot of the propagated error of normalized 2D scale-space derivatives (order 0-6). Each curve represents the ratio of variance of output to input noise of the linear normalized derivative of Gaussian filter vs. scale σ. Plot is on a log-log scale.

3.7. Discussion

Human vision is used as a benchmark for evaluating image analysis systems. It has long been known that humans are better able to distinguish different levels of contrast than absolute levels of intensity. Human vision is capable of distinguishing objects relative to their surroundings regardless of variations in lighting.

Scale space is often presented as a reasonable model for the overlapping receptive fields of the visual system. The results shown in this research demonstrate possible agreement between scale-space visual responses and the sensitivity of the human visual system to contrast rather than absolute intensity. I have analyzed the propagation of noise through scale-space intensity as well as scale-space derivative measurements. The results show that scale-space representations of first order, second order, and for 2D images, normalized scaled derivatives of up to the fifth order propagate less noise than the scale-space measures of absolute intensity.

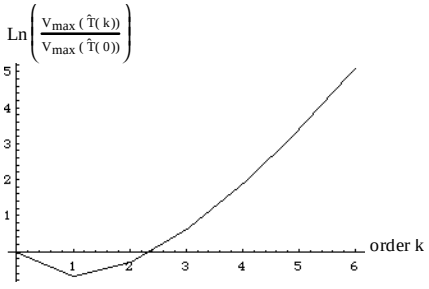

Figure 3.6. Plot of the propagated error of normalized 2D scale-space derivatives (order 0-6). Curve represents the ratio of variance of output to input noise of the linear unnormalized derivative of Gaussian filter vs. order of differentiation. Plot is on a log scale.

The implication is that a visual system based on scale-space gradients and scale-space curvature using the normalization suggested by Eberly should be more robust than a system based on absolute intensity. This implied agreement between scale-space analysis and the human visual system supplies circumstantial support for scale space as a visual model.

This finding is dependent on the form of normalization used to achieve dimensionless scale-space measurements. Eberly suggests multiplying the scale-space derivative measurement by σ^k, where σ is the aperture of the scale operation, and k is the order of differentiation. If $(\%_b)^k$ is substituted for the normalizing term (where b is some constant), dimensionless scale-space measurements are still achieved. However, this change in normalization changes the shape of the plots shown in figure 3.3. and figure 3.6. The global minimum represented in each of these plots can be altered to lie between any two values of k. This sensitivity to the propagation of noise in multiscale analysis to the selection of a normalization term requires further study.

3.8. Conclusion

I have derived analytic forms for the propagation of noise in an input signal through scale to calculated values for normalized scale-space derivatives. These expressions are based on the scale of the derivative-measuring kernel, the variance of the noise of the input signal, and the order of differentiation.

A comparison of the propagated error in both unnormalized and normalized derivatives is presented. The improved stability of derivative measurements as scale increases is clear. The assertion that the relative stability and accuracy of higher order derivatives of image intensity at large scale is shown to be incorrect when spatial measurements are made relative to the scale aperture.

An alternate finding is that for normalized scale-space derivatives, low order derivatives propagate less noise than either the zeroth order intensity measurement or derivatives of the third or higher orders. The first (gradient) and second (intensity curvature) spatial derivatives propagate less noise than the scale-space intensity value or derivatives of order four or higher. This finding is sensitive to the form of the normalization used to achieve dimensionless scale-space measurements. The normalization used in this study was suggested by Eberly. Dimensionless scale-space derivative measurements can be computed emphasizing the propagation of input noise to a lesser or greater degree. This sensitivity is of noise propagation to the normalization factor is still being explored.

Appendix : Covariance of Scale-Space Derivatives

The following derivation is a reconstruction of second central moments of multiscale 1D images with Gaussian additive white (spatially uncorrelated) noise. This work is adapted from Blom [Blom 1992]. For a full derivation of the general multidimensional case, see the appendix of Chapter 4 of Blom's dissertation.

To calculate the covariance of 1D scale-space derivatives, consider the relation in equation (3-22).

$$\mathrm{Cov}\left(L_{\tilde{I},x^i}(x\,|\,\sigma)\,,\,L_{\tilde{I},x^j}(x\,|\,\sigma)\right) = v_0 \int_{-\infty}^{\infty} \frac{\partial^i}{\partial x^i} G(\sigma,\tau)\,\frac{\partial^j}{\partial x^j} G(\sigma,\tau)\,d\tau \tag{3A-1}$$

Integrating (3A-1) by parts yields

$$\begin{aligned}\mathrm{Cov}\left(L_{\tilde{I},x^i}(x\,|\,\sigma)\,,\,L_{\tilde{I},x^j}(x\,|\,\sigma)\right) = v_0\left(\left(\frac{\partial^i}{\partial x^i} G(\sigma,x)\right)\left(\frac{\partial^{j-1}}{\partial x^{j-1}} G(\sigma,x)\right)\Big|_{-\infty}^{\infty}\right) \\ - v_0 \int_{-\infty}^{\infty} \frac{\partial^{i+1}}{\partial x^{i+1}} G(\sigma,\tau)\,\frac{\partial^{j-1}}{\partial x^{j-1}} G(\sigma,\tau)\,d\tau\end{aligned} \tag{3A-2}$$

Note that the first term of (3A-2) vanishes for large x. That is,

$$v_0\left(\left(\frac{\partial^i}{\partial x^i} G(\sigma,x)\right)\left(\frac{\partial^{j-1}}{\partial x^{j-1}} G(\sigma,x)\right)\Big|_{-\infty}^{\infty}\right) = 0 \tag{3A-3}$$

Repeating the integration by parts j times generates the following relationship

$$\mathrm{Cov}\left(L_{\tilde{I},x^i}(x\,|\,\sigma)\,,\,L_{\tilde{I},x^j}(x\,|\,\sigma)\right) = (-1)^j\, v_0 \int_{-\infty}^{\infty} \left(\frac{\partial^{i+j}}{\partial x^{i+j}} G(\sigma,x)\right) G(\sigma,x)\,dx \tag{3A-4}$$

If $w = \dfrac{x}{\sigma\sqrt{2}}$ and $dw = \dfrac{dx}{\sigma\sqrt{2}}$, then since $G(\sigma,x) = \dfrac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{x^2}{2\sigma^2}}$, (3A-4) becomes

$$\mathrm{Cov}\left(L_{\tilde{I},x^i}(x\,|\,\sigma)\,,\,L_{\tilde{I},x^j}(x\,|\,\sigma)\right) = (-1)^j\, v_0\, \frac{1}{2\pi\sigma^2}\left(\frac{1}{\sigma\sqrt{2}}\right)^{i+j} \int_{-\infty}^{\infty} \left(\frac{\partial^{i+j}}{\partial w^{i+j}} e^{-w^2}\right) e^{-w^2}\,dw \tag{3A-5}$$

Invoking Rodrigues' formula for Hermite polynomials:

$$\frac{\partial^k}{\partial z^k} e^{-z^2} = (-1)^k\, H_k(z)\, e^{-z^2} \tag{3A-6}$$

transforms equation (3A-5) into

$$\mathrm{Cov}\left(L_{\tilde{I},x^i}(x\,|\,\sigma)\,,\,L_{\tilde{I},x^j}(x\,|\,\sigma)\right) = (-1)^i\, v_0\, \frac{1}{2\pi\sigma^2}\left(\frac{1}{\sigma\sqrt{2}}\right)^{i+j} \int_{-\infty}^{\infty} H_{i+j}(w)\, e^{-2w^2}\,dw \tag{3A-7}$$

Consider only the integral element of (3A-7). Applying the recurrence relation of Hermite polynomials:

$$H_k(z) = 2z\, H_{k-1}(z) - 2(k-1)\, H_{k-2}(z) \quad , \qquad k \geq 2 \tag{3A-8}$$

transforms the integral element of (3A-7) to the following recurrence relation.

$$\int_{-\infty}^{\infty} H_{i+j}(w)\, e^{-2w^2} dw = \int_{-\infty}^{\infty} \left(2wH_{i+j-1}(w) - 2(i+j-1)H_{i+j-2}(w)\right) e^{-2w^2} dw \tag{3A-9}$$

simplifying to

$$\begin{aligned} &\int_{-\infty}^{\infty} H_{i+j}(w)\, e^{-2w^2} dw \\ &= \int_{-\infty}^{\infty} 2wH_{i+j-1}(w) e^{-2w^2} dw \\ &\quad - \int_{-\infty}^{\infty} 2(i+j-1)H_{i+j-2}(w)\, e^{-2w^2} dw \end{aligned} \tag{3A-10}$$

Integrating the first term of (3A-10) by parts yields

$$\begin{aligned} &\int_{-\infty}^{\infty} H_{i+j}(w)\, e^{-2w^2} dw \\ &= -\tfrac{1}{2}\left(H_{i+j-1}(w) e^{-2w^2} \Big|_{-\infty}^{\infty} \right) \\ &\quad + \int_{-\infty}^{\infty} \tfrac{1}{2} \tfrac{\partial}{\partial w} H_{i+j-1}(w) e^{-2w^2} dw \\ &\quad - \int_{-\infty}^{\infty} 2(i+j-1)H_{i+j-2}(w)\, e^{-2w^2} dw \end{aligned} \tag{3A-11}$$

Apply (3A-6), Rodrigues' formula for Hermite polynomials, to the first term of (3A-11).

$$\begin{aligned} &-\frac{1}{2} H_{i+j-1}(w) e^{-2w^2} \Big|_{-\infty}^{\infty} \\ &\quad = H_{i+j-1}(w) e^{-w^2} e^{-w^2} \Big|_{-\infty}^{\infty} \\ &\quad = (-1)^{i+j-1} e^{-w^2} \tfrac{\partial^{i+j-1}}{\partial w^{i+j-1}} e^{-w^2} \Big|_{-\infty}^{\infty} \\ &\quad = 0 \end{aligned} \tag{3A-12}$$

demonstrates that the first term vanishes for large values of w. Furthermore, combining (3A-6) and the recurrence relation shown in (3A-8) generates the following identity for Hermite polynomials:

$$\begin{aligned}
\tfrac{\partial}{\partial z} H_k(z) &= \tfrac{\partial}{\partial z}\left((-1)^k\, e^{z^2}\, \tfrac{\partial^k}{\partial z^k} e^{-z^2}\right) \\
&= 2z\left((-1)^k\, e^{z^2}\, \tfrac{\partial^k}{\partial z^k} e^{-z^2}\right) - \left((-1)^{k+1}\, e^{z^2}\, \tfrac{\partial^{k+1}}{\partial z^{k+1}} e^{-z^2}\right) \\
&= 2z\, H_k(z) - H_{k+1}(z) \\
&= 2z\, H_k(z) - \left(2z\, H_k(z) + 2k\, H_{k-1}(z)\right) \\
&= 2k\, H_{k-1}(z)
\end{aligned} \tag{3A-13}$$

Using the result of (3A-12) and applying equation (3A-13), equation (3A-11) becomes

$$\begin{aligned}
&\int_{-\infty}^{\infty} H_{i+j}(w)\, e^{-2w^2} dw \\
&= \int_{-\infty}^{\infty} (i + j - 1) H_{i+j-2}(w)\, e^{-2w^2} dw \\
&\quad - \int_{-\infty}^{\infty} 2(i + j - 1) H_{i+j-2}(w)\, e^{-2w^2} dw \\
&= -(i + j - 1) \int_{-\infty}^{\infty} H_{i+j-2}(w)\, e^{-2w^2} dw
\end{aligned} \tag{3A-14}$$

Summarizing the results shown between (3A-7) and (3A-14) reveals an important recurrence relation regarding covariances among 1D scale-space derivatives.

$$\int_{-\infty}^{\infty} H_{i+j}(w)\, e^{-2w^2} dw = -\left((i + j - 1) \int_{-\infty}^{\infty} H_{i+j-2}(w)\, e^{-2w^2} dw\right) \tag{3A-15}$$

Applying the recurrence relation in (3A-15) to (3A-7) yields

$$\begin{aligned}
&\mathrm{Cov}\left(L_{\tilde{I},x^i}(x \mid \sigma),\, L_{\tilde{I},x^j}(x \mid \sigma)\right) \\
&= (-1)^i\, v_0 \frac{1}{2\pi\sigma^2} \left(\frac{1}{\sigma\sqrt{2}}\right)^{i+j} \int_{-\infty}^{\infty} H_{i+j}(w)\, e^{-2y^2} dw \\
&= (-1)^{i+1}\, v_0 \frac{1}{2\pi\sigma^2} \left(\frac{1}{\sigma\sqrt{2}}\right)^{i+j} \left((i + j - 1) \int_{-\infty}^{\infty} H_{i+j-2}(w)\, e^{-2y^2} dw\right)
\end{aligned} \tag{3A-16}$$

Manipulating Rodrigues' formula of (3A-6), it is clear that $H_0(z) = 1$ and $H_1(z) = 2z$. If i+j is odd, then (i+j-1)/2 is an integer value. Repeated application of (3A-15) to (3A-7) through (i+j-1)/2 iterations yields

$$
\begin{aligned}
&\mathrm{Cov}\left(L_{\tilde{I},x^i}(x \mid \sigma)\,,\, L_{\tilde{I},x^j}(x \mid \sigma)\right) \\
&= (-1)^{i+\left(\frac{i+j-1}{2}\right)} \frac{v_0}{2\pi\sigma^2}\left(\frac{1}{\sigma\sqrt{2}}\right)^{i+j}\left(\prod_{r=0}^{(i+j-1)/2}(2r)\right)\left(\int_{-\infty}^{\infty} H_1(w)\, e^{-2w^2} dw\right) \\
&= (-1)^{i+\left(\frac{i+j-1}{2}\right)} \frac{v_0}{2\pi\sigma^2}\left(\frac{1}{\sigma\sqrt{2}}\right)^{i+j}\left(\prod_{r=0}^{(i+j-1)/2}(2r)\right)\left(\int_{-\infty}^{\infty} 2w\, e^{-2w^2} dw\right) \qquad ,\forall \text{ odd } i+j \qquad \text{(3A-17)} \\
&= (-1)^{i+\left(\frac{i+j-1}{2}\right)} \frac{v_0}{2\pi\sigma^2}\left(\frac{1}{\sigma\sqrt{2}}\right)^{i+j}\left(\prod_{r=0}^{(i+j-1)/2}(2r)\right)\left(-\frac{1}{2}\, e^{-2w^2}\Big|_{-\infty}^{\infty}\right) \\
&= 0
\end{aligned}
$$

Thus, for all odd i+j the corresponding covariance of the 1D scale-space derivatives is 0. This implies that odd derivatives are uncorrelated with even derivatives. If i+j is even, then (i+j)/2 is an integer value. Repeated application of (3A-7) to (3A-15) through (i+j)/2 iterations yields

$$
\begin{aligned}
&\mathrm{Cov}\left(L_{\tilde{I},x^i}(x \mid \sigma)\,,\, L_{\tilde{I},x^j}(x \mid \sigma)\right) \\
&= (-1)^{i+\left(\frac{i+j}{2}\right)} \frac{v_0}{2\pi\sigma^2}\left(\frac{1}{\sigma\sqrt{2}}\right)^{i+j}\left(\prod_{r=0}^{(i+j)/2}(2r-1)\right)\left(\int_{-\infty}^{\infty} H_0(w)\, e^{-2w^2} dw\right) \\
&= (-1)^{i+\left(\frac{i+j}{2}\right)} \frac{v_0}{2\pi\sigma^2}\left(\frac{1}{\sigma\sqrt{2}}\right)^{i+j}\left(\prod_{r=0}^{(i+j)/2}(2r-1)\right)\left(\int_{-\infty}^{\infty} e^{-2w^2} dw\right) \\
&= (-1)^{i+\left(\frac{i+j}{2}\right)} \frac{v_0}{2\sigma\sqrt{\pi}}\left(\frac{1}{\sigma\sqrt{2}}\right)^{i+j}\left(\prod_{r=0}^{(i+j)/2}(2r-1)\right)\left(\frac{1}{\sigma\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-\frac{x^2}{\sigma^2}} dx\right) \qquad ,\forall \text{ even } i+j \qquad \text{(3A-18)} \\
&= (-1)^{i+\left(\frac{i+j}{2}\right)} \frac{v_0}{2\sigma\sqrt{\pi}}\left(\frac{1}{\sigma\sqrt{2}}\right)^{i+j}\left(\prod_{r=0}^{(i+j)/2}(2r-1)\right)
\end{aligned}
$$

Finally, since variances are a special case (i = j) of the covariance calculation, generating an expression for the variance of a 1D scale-space derivative is straightforward.

$$
\begin{aligned}
V\left(L_{\tilde{I},k}(x \mid \sigma)\right) &= \mathrm{Cov}\left(L_{\tilde{I},k}(x \mid \sigma)\,,\, L_{\tilde{I},k}(x \mid \sigma)\right) \\
&= \frac{v_0}{2\sigma\sqrt{\pi}}\left(\frac{1}{\sigma^{2k}\, 2^k}\right)\left(\prod_{r=1}^{k}(2r-1)\right) \qquad \text{(3A-19)}
\end{aligned}
$$